Industry Self-Regulation of Consumer Data Privacy and Security
Siona Listokin
Associate Professor, School of Policy, Government and International Affairs
George Mason University
3351 Fairfax Drive, Arlington VA 22201
703-993-9756
slistoki@gmu.edu

Abstract
Industry self-regulation of consumer data privacy and security has been proposed as a flexible alternative and compliment to traditional government regulation. This study analyzes whether different types of existing industry-led standards improve online privacy and security. The paper examines which types of firms join voluntary standards and whether there is a difference in outcomes between trade association memberships (like the Digital Advertising Alliance) and certification programs (like TRUSTe). Results suggest that more trafficked websites are more likely to adopt standards, and that trade association membership does not have an effect on privacy and security performance. The article highlights the need for a valid privacy metric for robust empirical study of data privacy and security.

Introduction
As the creation and collection of consumer digital information continues its astonishing growth, consumer data privacy and security stand out as pressing areas of opportunity and concern for the online marketplace. In 2009 and 2012, the Federal Trade Commission (FTC) issued guidelines and best practices for self-regulation of consumer privacy and noted several industry-led initiatives covering the use of digital information that seek to foster innovation while protecting data privacy and security.\(^1\) Federal legislation proposals, such as the White

---

\(^1\) I am grateful for support from the George Mason Law School Law and Economics Center, as well as a research grant from Google, Inc. All ideas and errors are my own.

House’s Consumer Privacy Bill of Rights Act of 2015 relies heavily on “enforceable codes of conduct developed by diverse stakeholders,” that would surely include industry organizations.

There are a number of advantages to self-regulation in fast-changing industries like “e-commerce” (broadly defined) that collect and use consumer data. Information technology is fast-changing by nature and regulatory responses may not keep pace with the industry. When properly managed, self-regulation through trade associations and certification programs can adapt more quickly and appropriately to innovations than government regulation, and can provide a market solution to information asymmetries between firms and consumers by differentiating companies’ data privacy and security performance. Firms can use industry standards to increase consumer trust without stifling the creation of new products or shifting away from a free, advertising revenue-based business model. At the same time, self-regulation typically relies on self-policing for enforcement, creating conditions for adverse selection of firms that agree to comply with standards and moral hazard for firms once they are certified. Indeed, in May 2014, the FTC called for additional guidelines for “data brokers” to protect consumer privacy in part because monitoring and enforcement by self-regulating organizations have not sufficiently addressed regulators’ concerns.3

This study analyzes whether different types of existing industry-led standards for consumer data privacy and security effect online privacy. The paper examines which types of firms adopt voluntary standards, and whether there is a difference in outcomes between trade association memberships (like the Digital Advertising Alliance) and certification programs (like TRUSTe). Results suggest that more trafficked websites are more likely to be members of standards programs, and that trade association membership does not have an effect on privacy and security performance. While there is some evidence that certification can hurt subsequent privacy and security compared to similar non-certified websites, this result is sensitive to the privacy metric used.

This study contributes to a fairly sparse literature that evaluates the effects of membership in voluntary consumer data privacy standards. The paper also compares existing measures of online privacy and highlights the significant challenges in building a valid privacy metric.

Measures of website privacy that have been used in prior studies to evaluate the market for privacy frequently diverge and/or cover different websites, and empirical studies like this one may be highly sensitive to the choice of construct. The issues examined in this paper highlight the need for more rigorous empirical study and practical knowledge about different self-regulatory options in consumer data privacy.

Prior Literature

Consumer Data Privacy Self-Regulation

There are a number of studies that examine self-regulation in consumer data privacy. Scholars have introduced frameworks for viable self-regulation in this area. This literature considers self-regulation as a middle ground between a pure market model and government regulation, and includes industry trade associations and third party certification. Self-regulation, or industry regulation, can create rules, play a role in enforcement and/or be involved in adjudication. Analytical models consider consumers with heterogeneous preferences for privacy and noisy signals from firms as to privacy risk. Self-regulation (or seal of approval programs) can enhance trust in situations of isolated encounters, and may be a more efficient regime than mandatory regulation.

Empirical work in this area presents a mixed picture regarding the efficacy of industry regulation. A study of self-regulation via website privacy notices in 2000 noted that 1/3 of the sample websites did not post privacy policies and only 14% of those that did were comprehensive. Subsequent works in this area show an increase in the

---


5 See Swire, supra note 3.

6 See Tang et al, supra note 3.

existence of posted policies and variability in the content and
readability of these privacy notices.\textsuperscript{8} Other studies analyze specific
industry membership associations or certifications. Descriptive studies
that examine the content of the rules or enforcement mechanisms are
typically skeptical,\textsuperscript{9} though there are exceptions. For example,
Villafranco and Riley (2013) is generally positive about the industry
rules of NAI.\textsuperscript{10} A survey of consumers showed that only 11% of users
understand the NAI opt-out initiative.\textsuperscript{11}

I examine two different voluntary programs, TRUSTe (a third party
certification) and the DAA (an industry association). There are a few
notable studies that specifically test the efficacy of DAA membership or
TRUSTe certification beyond the case studies and review of standards
highlighted above. Komanduri et al (2011) examine the cookies
produced by the DAA and NAI opt-out mechanisms of the top 100
websites, and find numerous instances of non-compliance.\textsuperscript{12} In addition,
Edelman (2011) shows that TRUSTe certified websites are more likely
to be rated as untrustworthy, suggesting there is an adverse selection
effect for the certification seal.\textsuperscript{13} Miyazaki and Krishnamurthy (2002)

\textsuperscript{8} See, e.g., George R. Milne, Mary J. Culnan & Henry Greene, A
\textsuperscript{12} Komanduri et al, supra note 8.
\textsuperscript{13} Benjamin Edelman, Adverse Selection in Online “Trust” Certifications and Search Results, 10 ELECTRONIC COM. RES. & APPLICATIONS 17, 17-25 (2011).
compare sixty websites’ privacy policies and find that TRUSTe certification does not improve the content of privacy policies.¹⁴

While there are relatively few empirical evaluations of industry self-regulation of consumer data privacy, there are multiple studies of voluntary standards in other industries.¹⁵ Studies on self-regulation and third party certification in the financial, environmental, healthcare, food and other industries find mixed results as to the efficacy of self-regulation.¹⁶

Voluntary standards may serve as differentiating tools for consumers and regulators that indicate superior firm management and processes designed to protect data security or control privacy.¹⁷ Industry association membership or certification can improve certain outcomes through informal information exchange and industry pressure. Industry associations frequently facilitate communication between members and with regulators, and may host member conferences to disseminate research to firms. For example, the Digital Advertising Alliance calls for transparency about data collection, and clarifies the practical application of this principle through published cases on enhanced notice to consumers.¹⁸ “Mimetic forces” such as social networks and the

creation of guidelines and best practices can lead to compliance even in
the absence of sanctions.19

Hypothesis 1. Members of consumer data industry associations or
websites have better privacy and security than non-members.

While both industry associations and third-party certifications share
many characteristics conducive to effective private governance, there
are a number of key differences between these models. Trade
associations typically exist as self-organized institutions designed to
create collective governance structures and have incentives to establish
standards and behave as a “middle ground” between traditional
government regulation and the free market.20 Thus, industry
organizations are motivated to create standards that are sufficiently
restrictive to avoid external threats like government regulation and
enhance their members’ profiles with policymakers.

Certification organizations do not necessarily share this common goal
with their clients.21 In order for third party certifications or seals of
approval to convey a credible signal of quality, they must be
independent from those seeking the certification. There is little near-
term incentive for for-profit certifiers to restrict their membership.22 In
addition, certified websites are not integral to – and may not even be
involved in – the creation of certification standards as they are in
industry associations. It is plausible that within the same industry,
trade or industry-led associations may have different impacts than
certifications.

Hypothesis 2. Paid privacy certifications (seals) do not improve privacy
performance.

Data
My sample includes the top 10,000 trafficked websites in 2015 that
existed in 2007 and 2010. I lose about 100 websites that did not have
industry information. Summary statistics are shown in Table 1. Website
traffic is an imperfect screening technique for the population of
comparison sites for DAA membership or TRUSTe certification. For

19 See King & Lenox, supra note 15.
20 See ELINOR OSTROM, GOVERNING THE COMMONS: THE EVOLUTION OF
INSTITUTIONS FOR COLLECTIVE ACTION (1990).
21 See Miyazaki & Krishnamurthy, supra note 13.
22 See Bob Tanner, Independent Assessment by Third-Party Certification
Bodies, 11 FOOD CONTROL 415, 415-17 (2000).
example, the DAA includes a number of advertising network companies that do not draw much traffic. Thus, only 1% of the websites in my sample are members of the DAA in 2015, and about 4% of the websites are TRUSTe certified despite larger member rolls.23

**DAA and TRUSTe**

TRUSTe is a private data privacy management company, with thousands of certified firms as customers. TRUSTe sells software that provides a data privacy management platform and provides assessments and certification for over 4,000 firms that meet the program requirements. The certification is primarily concerned with transparency and consumer choice, and includes the requirement that firms implement commercially reasonable protections for data security.24

It is worth noting that one of the TRUSTe programs ("TRUSTed Data") bases its requirements in part on the FTC self-regulatory principles, the NAI principles and the DAA principles.25 While the TRUSTe certification provides specific monitoring guidelines, it is not clear how it deals with non-compliance. The seal requires annual recertification, and the consumer dispute resolution service processes thousands of consumer complaints every year. In 2014, the FTC charged that TRUSTe failed to follow its guidelines for privacy seal recertifications in over 1,000 incidences between 2006 and 2013; TRUSTe eventually settled with the FTC.26

The Digital Advertising Alliance (DAA) is a non-profit organization made up of marketing and advertising industry associations that seeks to provide self-regulatory consumer privacy principles for internet based advertising. The DAA is one of the most prominent self-regulation associations in consumer data privacy and security, but has been criticized for promoting weak data privacy programs and enforcement.27

---

23 The sample is also screened for websites that exclusively carry adult content.
24 https://www.truste.com/privacy-certification-standards/
27 https://www.cylab.cmu.edu/files/pdfs/tech_reports/CMUCyLab12008.pdf
The DAA was established in 2009 by several U.S. advertising associations, following the release of a Federal Trade Commission (FTC) report on “Self-Regulatory Principles for Online Behavioral Advertising.” It is led by the Association of National Advertisers, The American Advertising Federation, 4A’s, Network Advertising Initiative, Direct Marketing Association and Interactive Advertising Bureau. Originally, participating companies consisted of advertisers and third party analytics companies, but starting in 2011, DAA expanded its efforts to include social networks and non-advertising firms. Enforcement is handled by the Direct Marketing Association (DMA) and the Council for Better Business Bureau (CBBB); DAA refers to their “independent enforcement component” though it is worth noting that these organizations are participating and founding associations.

Privacy Metrics
The choice of privacy metric is crucial to this study and other articles that attempt to evaluate website privacy and security. While researching website privacy and security performance, I came across a number of potential measures, including some that have been used in other research. Some of these measures are included in the summary statistics and briefly described in Table 2.

I use two measures of privacy and security, TrustGauge and Privacy Rights Clearinghouse data breach records. The strengths and weaknesses of these measures are discussed below.

TrustGauge is an index that measures a 10-point “trust score” based on a number of different website factors. The index is not meant to exclusively or comprehensively measure privacy; it is more a construct of trustworthiness or validity. Nonetheless, the measure is based in part on website privacy policies and security. The first set of features is focused on website content and verifiability, and scores sites based on the availability of contact information, privacy statements and verified

28 2009 FTC STAFF REPORT, supra note 1.
29 http://www.digitaladvertisingalliance.org/content.aspx?page=enforcement
30 Of course, many potential measures are not included in this table. Notable absences include the Web of Trust (WOT), MSCI Privacy Index, EFF’s Who Has Your Back?, and Terms of Service; Didn’t Read. These alternatives either don’t cover enough websites, are relatively new and unknown, or are very similar to the metrics included in the paper.
31 http://www.trustgauge.com/about.html
customer service response. The second group of factors concerns security, such as using secure protocols on billing pages. A final set of features measures third party certification and website traffic. I subtract the points allotted for TRUSTe certification for any certified websites (no points are added for DAA membership) but do not adjust scores for website rank. As a result, I expect larger websites to have higher scores by design, and I must control for the traffic rank. I use a dummy that equals one for websites that score above 5 (the midway point). I have two full years of TrustGauge data (2007 and 2015) for the top 10,000 ranked websites in either year and a partial sample for 2010. In 2015, only 6% of the websites in my sample are coded as untrustworthy (down from about 11% in 2007).

Privacy Rights Clearinghouse (PRC) is a California nonprofit organization that has collected information about reported data breaches since 2005.32 Data breaches are a particular element of data privacy and security. A reported breach, or absence of a breach is not necessarily an indication of underlying security weakness. In addition, the breaches reported appear to be heavily concentrated among the most trafficked websites and firms. Most breaches included in their data involve social security or account numbers and other sensitive information. The PRC collects the total number of records that were compromised in each breach, however this number is “unknown” for many of the observations. In addition, the number of breached records may not be a proper indication of lack of data security. For example, a substantial breach in Sony Pictures Entertainment in late 2014 records only 47,000 compromised records (as compared to 101.6 million records in a 2011 breach of the PlayStation Network and Sony Online Entertainment). Unfortunately, these data are not consistently reported, and I create a measure based solely on the existence of a reported breach. I restrict the PRC data to non-governmental or educational targets and create a cumulative dummy each year for having ever been breached (i.e. the 2015 dummy will be 1 if the website has had at least one breach since 2005). Less than one percent of websites report more than one breach over the time period. About 2.5% of the websites in my sample have had a reported data breach by 2015.

Considering that neither TrustGauge or PRC were created as measures of data privacy or security, it is worth comparing these constructs to other existing indices. I test the correlations of TrustGauge and the data

---

32 https://www.privacyrights.org/data-breach
breach records as a check of their legitimacy. The correlation table of privacy measures in 2015 is shown in Table 3. Frankly, the results are discouraging, and potentially consequential for the broader research area that evaluates data privacy. Pairwise spearman correlations are fairly low (and sometimes negative) and generally not significant. TrustGauge and breach records do not have a particularly strong (or weak) correlation to other measures. I use these metrics because of the availability of past data. The inconsistencies between and lack of convergence among privacy measures is a major limitation of empirical study in this area. While other measures may prove to be stronger with time, TrustGauge and PRC have a relatively long history and cover a range of websites and companies.

I also collect information on website traffic rank and industry, from a well-known web traffic site, ranking.com.

**Results**

My empirical strategy is to test differences in privacy outcomes for member/certified sites over time, and to refine the validity of these comparisons to approach the counterfactual: how would these websites perform without membership or certification? I begin by comparing average privacy performance of members to non-members, and then restrict the sample to the most trafficked websites in order to control for the effects of site size. Finally, I match certified websites to uncertified website and employ a difference-in-differences approach to help differentiate between selection and treatment effects of membership/certification.

Initial t-tests of privacy and security show that both DAA members and TRUSTe certified websites are more likely to be rated highly by TrustGauge, and that this difference is significant (Table 4). This result supports hypothesis 1. At the same time, these websites are also more likely to have experienced a data breach.

In order to examine whether these differences persist over time and across different types of websites, I test the difference in privacy performance in both 2007 and 2015. These results are shown graphically in Figures 1 and 2 for DAA membership and TRUSTe certification, respectively. Column 1 in both figures show the likelihood of being well-rated by TrustGauge and of having a reported data breach in both time periods along with 95% confidence intervals, for the full sample of websites in the study. For the most part, the mixed results suggested by the initial ttests persist. Member/certified sites are significantly more likely to be rated trustworthy in 2007, though this
gap shrinks by 2015. These same websites are also more likely to have been breached by 2007; this difference actually grows larger by 2015. However, further tests suggest that many of these differences are related to website traffic. Column 2 of Figures 1 and 2 show the same tests over time for only the top 500 trafficked websites. The results are quite different. DAA members are essentially indistinguishable from non-members in terms of TrustGauge and breaches. TRUSTe certification is similarly not a differentiator by 2015, though the 2007 gaps remain.

To better control for the effects of size and other factors, I use a difference-in-differences approach. I match member/certified websites to a control group of non-certified sites and compare the difference in privacy outcomes in 2015 and 2007 (the difference between certified and non-certified websites in 2015 minus the difference between certified and non-certified websites in 2007). The matching step is intended to provide the counterfactual for how websites would have performed if they had not been certified, by eliminating time-constant unobserved effects on outcomes. I use propensity score matching to create better control websites that have the same probability of being certified based on website traffic rank and industry in 2007 (Rosenbaum and Rubin 1983).

The difference-in-differences approach typically runs across a pre- and post-treatment period. Since both DAA membership and TRUSTe certification pre-date 2007, however, I do not have a strict “pre-treatment” observation. I therefore run two tests with different treatment groups. In model A, treated websites are those that were members/certified in both 2007 and 2015. “Consistent” membership cuts the number of certified sites for both DAA and TRUSTe, to 25 and 58 websites, respectively. Model B models a pre- and post-treatment by

34 The set of firms that are certified or members differs between 2007 and 2015. Websites both join and dropout of certification. The test results shown in figures 1 and 2 remain almost exactly the same if I restrict the sample to websites that are consistent members, or if only compare sites that dropout. Given the subsequent results shown in this section, this is further confirmation that certification – at any point – is correlated with other factors that impact privacy and security.

restricting certification to sites to those that were not initially certified in 2007, but were by 2015.

Figure 3 shows the propensity score for model A certified and non-certified sites before and after matching for both DAA and TRUSTe. Results are similar for model B. In both cases, the certified sites are quite different insofar as propensity scores than the full sample of non-certified sites; the match improves the comparison considerably.

The difference-in-differences results are shown in Table 5. The first four columns show the results for DAA membership, using TrustGauge and breaches as the outcome variables for both models A and B. The certification dummy shows the difference in TrustGauge/breach averages in 2015 minus the difference in TrustGauge/breach in 2007. For the most part, the results are not significant, suggesting that trade association membership does not impact privacy. TRUSTe has a negative impact on privacy and security, though these results are not consistent across all models. Subsequent to TRUSTe certification, websites are almost 5% less likely to be rated to trustworthy compared to the trend among non-certified sites (column 6). However, this result only holds for model B, suggesting that there is an unobserved difference between websites that were certified in 2007 to those that subsequently became certified. Certification also increases the likelihood of having been breached by 2015 compared to the trend of the control websites (columns 7 and 8). These results support hypothesis 2, that paid certifications do not improve privacy.

Discussion
This study evaluates industry self-regulation and paid certification of data privacy and security. Overall, the results suggest that the populations of websites that join industry associations or get certified are different than those that do not: they are simultaneously more trustworthy and more likely to have had a reported data breach. However, industry association membership does not appear to improve the likelihood of being trustworthy or secure. I find some evidence that the paid certification seal TRUSTe actually hurts performance compared to similar non-certified sites, as measured by TrustGauge and PRC. These results extend a similar study that finds TRUSTe certified sites are less trustworthy.36

---

36 Edelman, supra note 12.
As discussed previously, the privacy and security metrics used in this study are imperfect, and are not available for a true longitudinal study. Identifying an accurate measure of data privacy and security performance is difficult, and this work demonstrates that many existing metrics diverge and cover different populations of websites. The lack of convergence in the measures may undermine implications I would draw from this current study. More importantly, the inconsistency of privacy metrics has implications for the “economics of privacy” field. Empirical tests of privacy markets frequently require independent assessments of website privacy, and the construct validity of existing metrics is suspect. This is a developing area, and new measures are still being created. At the very least, the strength of privacy measures is an area ripe for future research.

With the above caveat in mind, the current results have implications for websites, policymakers and users. Industry associations like the DAA are meant to generate a set of best practices and to communicate with policymakers. DAA membership does not have a discernible impact on privacy performance as measured by TrustGauge. It is also possible that the DAA does improve privacy performance or the regulatory environment, but these benefits are not exclusive to members and thus fail to differentiate members from non-members. DAA could therefore facilitate industry self-regulation without confining those results to its small membership group. Paid certifications like TRUSTe may provide process management benefits, but there is no evidence in this study that its seals improve outcomes (in fact, outcomes appear to suffer). This result is in line with the FTC’s complaint that TRUSTe did not perform recertifications as promised for years.

In general, the study highlights the relative dearth of empirical evaluations of certification programs and self-regulation of data privacy. Further research is needed to create a more nuanced picture of website privacy and security performance. As the public, government, and FTC specifically, continue to consider data privacy as a product subject to contract, transaction and fair treatment the need for practical assessments of privacy will grow. This study outlines a research strategy to evaluate website privacy and certifications across different contexts going forward.

References


<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Certification/Membership</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRUSTe (2007)</td>
<td>9904</td>
<td>0.031</td>
<td>0.174</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>TRUSTe (2015)</td>
<td>9904</td>
<td>0.037</td>
<td>0.188</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>DAA (2007)</td>
<td>9904</td>
<td>0.008</td>
<td>0.091</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>DAA (2015)</td>
<td>9904</td>
<td>0.011</td>
<td>0.105</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Privacy Metrics (outcome var's)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TrustGauge</td>
<td>9904</td>
<td>0.945</td>
<td>0.227</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ever Breached</td>
<td>9904</td>
<td>0.026</td>
<td>0.160</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Other Privacy Metrics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Privacy Grade Scale</td>
<td>974</td>
<td>3.723</td>
<td>0.690</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Disconnect Expected Use</td>
<td>836</td>
<td>-0.083</td>
<td>0.466</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>Disconnect Expected Collection</td>
<td>836</td>
<td>-0.254</td>
<td>0.592</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>Disconnect Location Data</td>
<td>186</td>
<td>-0.667</td>
<td>0.566</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>Disconnect Data Retention</td>
<td>836</td>
<td>-0.543</td>
<td>0.506</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>Disconnect SSL Support</td>
<td>9904</td>
<td>-0.370</td>
<td>0.929</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>Disconnect Heartbleed Fix</td>
<td>3116</td>
<td>0.688</td>
<td>0.478</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>SiteAdvisor</td>
<td>9904</td>
<td>0.894</td>
<td>0.308</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

2015 data unless otherwise noted
<table>
<thead>
<tr>
<th>Name</th>
<th>Intent</th>
<th>Population</th>
<th>Industry Curve?</th>
</tr>
</thead>
<tbody>
<tr>
<td>TrustGauge</td>
<td>TrustGauge scores websites based on their trustworthiness. Scores are based on the existence of legitimate contact information, privacy policies, secure payments, certification and traffic</td>
<td>1,000,000+ websites</td>
<td>N</td>
</tr>
<tr>
<td>Privacy Rights Clearinghouse</td>
<td>PRC collects information about reported data breaches that has compromised personal information. The information is pulled from a number of sources including the Open Security Foundation and the California Attorney General.</td>
<td>Reported data breaches since 2005</td>
<td>NA</td>
</tr>
<tr>
<td>Privacy Grade</td>
<td>Privacy Grade is created by a group of researchers at Carnegie Mellon's CHIMPS lab. It grades mobile apps based on user's privacy expectations</td>
<td>Over 50,000 free smartphone apps</td>
<td>Y</td>
</tr>
<tr>
<td>Disconnect Privacy Icons</td>
<td>Disconnect.me Privacy Icons shows a set of icons that summarizes important elements of privacy policies and security practices. The icons cover areas like expected use and collection of data, the use of location data, data retention policies, SSL support and Heartbleed fixes. Privacy Icons pulls information from TRUSTe's Privacy Policy Database, which is separate from their certification program</td>
<td>Over 5,000 websites</td>
<td>N</td>
</tr>
<tr>
<td>SiteAdvisor</td>
<td>SiteAdvisor rates websites based on their security by assessing email spam and spyware.</td>
<td>95% of the Web</td>
<td>N</td>
</tr>
</tbody>
</table>
Table 3. Privacy metric correlations (pairwise spearman correlations 2015)

<table>
<thead>
<tr>
<th></th>
<th>TrustGauge</th>
<th>Ever Breached</th>
<th>Privacy Grade Scale</th>
<th>Privacy Grade Dummy</th>
<th>Disconnect Expected Use</th>
<th>Disconnect Expected Collection</th>
<th>Disconnect Location Data</th>
<th>Disconnect Data Retention</th>
<th>Disconnect SSL Support</th>
<th>Disconnect Heartbleed Fix</th>
<th>SiteAdvisor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TrustGauge</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ever Breached</strong></td>
<td>0.0688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Privacy Grade Scale</strong></td>
<td>-0.0455</td>
<td>0.1553</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Privacy Grade Dummy</strong></td>
<td>-0.0606</td>
<td>0.1282</td>
<td>0.9583*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disconnect Expected Use</strong></td>
<td>-0.0229</td>
<td>0.1</td>
<td>0.0293</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disconnect Expected Collection</strong></td>
<td>0.0876*</td>
<td>-0.1162</td>
<td>-0.1375*</td>
<td>-0.1438*</td>
<td>0.2516*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disconnect Location Data</strong></td>
<td>0.0882</td>
<td>0.0248</td>
<td>-0.1296</td>
<td>-0.1029</td>
<td>0.1987*</td>
<td>0.1452</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disconnect Data Retention</strong></td>
<td>-0.0318</td>
<td>-0.1806</td>
<td>-0.0235</td>
<td>-0.01</td>
<td>0.0289</td>
<td>0.0186</td>
<td>0.3544</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disconnect SSL Support</strong></td>
<td>-0.0287*</td>
<td>-0.0452</td>
<td>0.0261</td>
<td>0.01</td>
<td>-0.0174</td>
<td>-0.0283</td>
<td>0.0842</td>
<td>0.0615</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Disconnect Heartbleed Fix</strong></td>
<td>-0.206*</td>
<td>-0.2261*</td>
<td>0.0284</td>
<td>0.0622</td>
<td>-0.0267</td>
<td>-0.0875</td>
<td>-0.1073</td>
<td>-0.0738</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SiteAdvisor</strong></td>
<td>0.2098*</td>
<td>0.1098</td>
<td>-0.0864*</td>
<td>-0.0863*</td>
<td>0.0427</td>
<td>0.039</td>
<td>0.064</td>
<td>0.0898*</td>
<td>0.019</td>
<td>-0.1488*</td>
<td></td>
</tr>
</tbody>
</table>

*p<0.05
Privacy Grade covers a larger population of smartphone applications. I matched websites to identifiable apps for the top 5,000 websites.
### Table 4. Ttests

#### DAA

<table>
<thead>
<tr>
<th>2015 Membership Status</th>
<th>N</th>
<th>TrustGauge</th>
<th>Data Breach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Member</td>
<td>110</td>
<td>0.982</td>
<td>0.209</td>
</tr>
<tr>
<td>Non-member</td>
<td>9794</td>
<td>0.893</td>
<td>0.024</td>
</tr>
<tr>
<td>Difference</td>
<td>9904</td>
<td>0.089***</td>
<td>0.185***</td>
</tr>
<tr>
<td>Std Error</td>
<td></td>
<td>0.029</td>
<td>0.015</td>
</tr>
</tbody>
</table>

#### TRUSTe

<table>
<thead>
<tr>
<th>2015 Certification Status</th>
<th>N</th>
<th>TrustGauge</th>
<th>Data Breach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Certified</td>
<td>309</td>
<td>0.984</td>
<td>0.081</td>
</tr>
<tr>
<td>Non-certified</td>
<td>9595</td>
<td>0.891</td>
<td>0.024</td>
</tr>
<tr>
<td>Difference</td>
<td>9904</td>
<td>0.093***</td>
<td>0.057***</td>
</tr>
<tr>
<td>Std Error</td>
<td></td>
<td>0.018</td>
<td>0.009</td>
</tr>
</tbody>
</table>

*p < 0.10, **p < 0.05, ***p < 0.01
Table 5. Difference in Difference Estimates

<table>
<thead>
<tr>
<th>Certification Type</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DAA</td>
<td></td>
<td></td>
<td></td>
<td>TRUSTe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outcome Variable</td>
<td>TrustGauge</td>
<td>TrustGauge</td>
<td>Ever Breached</td>
<td>Ever Breached</td>
<td>TrustGauge</td>
<td>TrustGauge</td>
<td>Ever Breached</td>
<td>Ever Breached</td>
</tr>
<tr>
<td>Certification (Dummy)</td>
<td>-0.022</td>
<td>-0.017</td>
<td>0.055</td>
<td>0.025</td>
<td>-0.303</td>
<td>-0.049***</td>
<td>0.088***</td>
<td>0.039***</td>
</tr>
<tr>
<td>SE</td>
<td>0.033</td>
<td>0.034</td>
<td>0.051</td>
<td>0.038</td>
<td>0.021</td>
<td>0.022</td>
<td>0.021</td>
<td>0.017</td>
</tr>
<tr>
<td>N</td>
<td>9704</td>
<td>9704</td>
<td>9704</td>
<td>9704</td>
<td>9704</td>
<td>9704</td>
<td>9704</td>
<td>9704</td>
</tr>
<tr>
<td>Certified Websites</td>
<td>25</td>
<td>93</td>
<td>25</td>
<td>93</td>
<td>58</td>
<td>268</td>
<td>58</td>
<td>268</td>
</tr>
</tbody>
</table>

Two different treatment groups are used. In Model A, the treated websites were members/certified in 2007 (2009) and 2015. In Model B, the treated websites are only certified in 2015.

The Certification dummy is the difference-in-difference estimate. It measures the difference between treated and control websites differences in 2015 and treated and control websites differences in 2007.

Websites are matched on industry and web traffic rank.

*p < 0.10, **p < 0.05, ***p < 0.01
Figure 1.
DAA Membership vs Non-Membership

Column 1

Column 2
Figure 2.
TRUSTe Certification vs. Non-certification
Figure 3. Propensity score before and after match

DAA

TRUSTe